

Tue May 15 14:46:35 2001

us-09-373-230-5.rst

Page 23

Db 13 TTTGAGAGATGAGCC 28

Search completed: May 15, 2001, 10:54:39

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2001, 10:05:34 ; Search time 1131.41 Seconds
(without alignments)
131.264 Million cell updates/sec

Title: US-09-373-230-5

Perfect score: 17

Sequence: 1 TTYGAGAGATGAGCC 17

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9623517 segs, 4368049070 residues
Total number of hits satisfying chosen parameters: 19247034

Minimum DB seg length: 0
Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: em_est33:*
34: em_est34:*
35: em_est35:*
36: em_est36:*

37: em_est37:*
38: em_est38:*
39: em_est39:*
40: em_est40:*
41: em_est41:*
42: em_est42:*
43: em_est43:*
44: em_est44:*
45: em_est45:*
46: em_est46:*
47: em_est47:*
48: em_est48:*
49: em_est49:*
50: em_est50:*
51: em_est51:*
52: em_est52:*
53: em_est53:*
54: em_est54:*
55: em_est55:*
56: em_est56:*
57: em_est57:*
58: em_est58:*
59: em_est59:*
60: em_est60:*
61: em_est61:*
62: em_est62:*
63: em_est63:*
64: em_est64:*
65: em_est65:*
66: em_est66:*
67: em_est67:*
68: em_est68:*
69: em_est69:*
70: em_est70:*
71: em_est71:*
72: em_est72:*
73: em_est73:*
74: em_est74:*
75: em_est75:*
76: em_est76:*
77: em_est77:*
78: em_est78:*
79: em_est79:*
80: em_est80:*
81: em_est81:*
82: em_est82:*
83: em_est83:*
84: em_est84:*
85: em_est85:*
86: em_est86:*
87: em_est87:*
88: em_est88:*
89: em_est89:*
90: em_est90:*
91: em_est91:*
92: em_est92:*
93: em_est93:*
94: em_est94:*
95: em_est95:*
96: em_est96:*
97: em_est97:*
98: em_est98:*
99: em_est99:*
100: em_est100:*
101: em_est101:*
102: em_est102:*
103: em_est103:*
104: em_est104:*
105: em_est105:*
106: em_est106:*
107: em_est107:*
108: em_est108:*
109: em_est109:*
110: em_est110:*
111: em_est111:*
112: em_est112:*
113: em_est113:*
114: em_est114:*
115: em_est115:*
116: em_est116:*
117: em_est117:*
118: em_est118:*
119: em_est119:*
120: em_est120:*
121: em_est121:*
122: em_est122:*
123: em_est123:*
124: em_est124:*
125: em_est125:*
126: em_est126:*
127: em_est127:*
128: em_est128:*
129: em_est129:*
130: em_est130:*
131: em_est131:*
132: em_est132:*

110: gb_est41:*
 111: gb_est42:*
 112: gb_est43:*
 113: gb_est44:*
 114: gb_est45:*
 115: gb_est46:*
 116: gb_est47:*
 117: gb_est48:*
 118: gb_est49:*
 119: gb_est50:*
 120: gb_est51:*
 121: gb_est52:*
 122: gb_est53:*
 123: gb_est54:*
 124: gb_est55:*
 125: gb_est56:*
 126: gb_est57:*
 127: gb_est58:*
 128: gb_est59:*
 129: gb_est60:*
 130: gb_est61:*
 131: gb_est62:*
 132: gb_est63:*
 133: gb_est64:*
 134: gb_est65:*
 135: gb_est66:*
 136: gb_est67:*
 137: gb_est68:*
 138: gb_est69:*
 139: gb_est70:*
 140: gb_est71:*
 141: gb_est72:*
 142: gb_est73:*
 143: gb_est74:*
 144: gb_est75:*
 145: gb_est76:*
 146: gb_est77:*
 147: gb_est78:*
 148: gb_est79:*
 149: gb_est80:*
 150: gb_est81:*
 151: gb_est82:*
 152: gb_est83:*
 153: gb_est84:*
 154: gb_est85:*
 155: gb_est86:*
 156: gb_est87:*
 157: gb_est88:*
 158: gb_est89:*
 159: gb_est90:*
 160: gb_est91:*
 161: gb_est92:*
 162: gb_est93:*
 163: gb_est94:*
 164: gb_est95:*
 165: gb_est96:*
 166: gb_est97:*
 167: gb_est98:*
 168: gb_est99:*
 169: gb_est100:*
 170: gb_est101:*
 171: gb_est102:*
 172: gb_est103:*
 173: gb_est104:*
 174: gb_est105:*
 175: gb_est106:*
 176: gb_est107:*
 177: gb_est108:*
 178: gb_est109:*
 179: gb_est110:*
 180: gb_est111:*
 181: gb_est112:*
 182: gb_est113:*

183: em_gss_hum7:*
 184: em_gss_hum8:*
 185: em_gss_hum9:*
 186: em_gss_inv1:*
 187: em_gss_inv2:*
 188: em_gss_inv3:*
 189: em_gss_inv4:*
 190: em_gss_inv5:*
 191: em_gss_inv6:*
 192: em_gss_inv7:*
 193: em_gss_inv8:*
 194: em_gss_inv9:*
 195: em_gss_inv10:*
 196: em_gss_inv11:*
 197: em_gss_inv12:*
 198: em_gss_inv13:*
 199: em_gss_inv14:*
 200: em_gss_inv15:*
 201: em_gss_inv16:*
 202: em_gss_inv17:*
 203: em_gss_inv18:*
 204: em_gss_inv19:*
 205: em_gss_inv20:*
 206: em_gss_inv21:*
 207: em_gss_inv22:*
 208: em_gss_inv23:*
 209: em_gss_inv24:*
 210: em_gss_inv25:*
 211: em_gss_inv26:*
 212: em_gss_inv27:*
 213: em_gss_inv28:*
 214: em_gss_inv29:*
 215: em_gss_inv30:*
 216: em_gss_inv31:*
 217: em_gss_inv32:*
 218: em_gss_inv33:*
 219: em_gss_inv34:*
 220: em_gss_inv35:*
 221: em_gss_inv36:*
 222: em_gss_inv37:*
 223: em_gss_inv38:*
 224: em_gss_inv39:*
 225: em_gss_inv40:*
 226: em_gss_inv41:*
 227: em_gss_inv42:*
 228: em_gss_inv43:*
 229: em_gss_inv44:*
 230: em_gss_inv45:*
 231: em_gss_inv46:*
 232: em_gss_inv47:*
 233: em_gss_inv48:*
 234: em_gss_inv49:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	15.4	90.6	143	AA659492	AA659492 nu25005.s
2	15.4	90.6	264	BB318243	BB318243 BB318243
3	15.4	90.6	327	AA054913	AA054913 SRAWCA150
4	15.4	90.6	335	T01405	T01405 WEST02126 E
5	15.4	90.6	375	AA673096	AA673096 ba57C02.y
6	15.4	90.6	434	AA237736	AA237736 mx2901.r
7	15.4	90.6	444	AA505386	AA505386 UI-HF-BNO
8	15.4	90.6	460	BF660764	BF660764 maat2e04.
9	15.4	90.6	474	AI121020	AI121020 ud69C02.x
10	15.4	90.6	499	BG076302	BG076302 H3158B03-
11	15.4	90.6	504	BF453847	BF453847 maat2e04.

C	12	15.4	90.6	512	111	AM125102
C	13	15.4	90.6	564	173	BC088827
C	14	15.4	90.6	603	13	AA930362
C	15	15.4	90.6	608	12	AA842840
C	16	15.4	90.6	713	142	BE979995
C	17	15.4	90.6	719	136	BE544528
C	18	15.4	90.6	740	229	AG007422
C	19	15.4	90.6	772	169	BF793924
C	20	15.4	90.6	779	136	BE542842
C	21	15.4	90.6	872	223	AZ530862
C	22	15.4	90.6	921	231	CNS0447P
C	23	15.4	90.6	937	136	BE544857
C	24	15.4	84.7	252	119	AM741070
C	25	14.4	84.7	267	26	AV261563
C	26	14.4	84.7	279	121	AM930594
C	27	14.4	84.7	306	116	AM487887
C	28	14.4	84.7	325	114	AM356632
C	29	14.4	84.7	330	159	243844
C	30	14.4	84.7	342	148	BF461899
C	31	14.4	84.7	343	13	AA885264
C	32	14.4	84.7	351	102	AI836656
C	33	14.4	84.7	353	222	AZ498397
C	34	14.4	84.7	350	175	D37330
C	35	14.4	84.7	377	175	C45493
C	36	14.4	84.7	380	32	AM028991
C	37	14.4	84.7	386	103	AI866547
C	38	14.4	84.7	398	156	T09111
C	39	14.4	84.7	423	116	AM450430
C	40	14.4	84.7	423	2	AA107198
C	41	14.4	84.7	428	21	AI539998
C	42	14.4	84.7	429	2	AA140177
C	43	14.4	84.7	429	206	AQ434635
C	44	14.4	84.7	433	219	AZ318553
C	45	14.4	84.7	457	11	AA770992

AM125102	UI-M-BH2
BC088827	H1-15B03-
AA930362	V95F107.r
AA842840	SWANCA03
BE979995	UI-M-BG2-
BE544528	601.07B354
AG007422	Homo sapi
BF793924	602545534
BE542842	601067660
AZ530862	ENRAC19PR
CNS0447P	601.07B856
BE544857	Genotox
AM741070	uq94q12.y
AV261563	EST31583
AM930594	EST356437
AM487887	UI-M-BH3-
AM356632	39122 MAR
243844	HSC1MB11
BF461899	UI-M-CGDP
AA885264	am40d12.s
AI836656	UI-M-AP0-
AZ498397	130335K07
D37330	CELOK044XF
C45493	C45493
AM028991	EST733246
AI866547	4250602.x
T09111	EST07004
AM450430	UI-M-BH3-
AA107198	ml18605.r
AI539998	SMOVARCAP
AA140177	mb93f03.r
AQ434635	HS_5086.B
AZ318553	IM007J19
AA770992	vt14e05.r

ALIGNMENTS

RESULT 1
LOCUS AA659492/c 143 bp mRNA EST 05-NOV-1997
DEFINITION nu2305.s1 NCI_CGAP_Pri2 Homo sapiens cDNA clone IMAGE:1209033, mRNA
ACCESSION AA659492
VERSION AA659492.1 GI:2595646
KEYWORDS EST.
SOURCE human
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 143)
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

JOURNAL COMMENT
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuqui, M.D.,
Michael Emmert-Buck, M.D., Ph.D.,
CDNA Library Preparation: David B. Krizman, Ph.D.,
DNA Sequencing by: Genome Systems Inc., Greg Lennon, Ph.D.
Clone distribution: NCI-CGAP clone distribution Information Center
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/bdip/image/image.html
Seq primer: -40m3 fwd. ET from Amersham
High quality sequence stop: 139.
Location/Qualifiers
1. 143
/organism="Homo sapiens"
/db_xref="taxon:9606"

/clone="IMAGE:1209033"
/clone_id="NCI_CGAP_Pri2"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
/note="Vector: pAMP10; Site 1: NciI; Site 2: EcoRI; 1st
strand cDNA was primed with oligo(dT)17 on 50 ng of
DNase-treated, total cellular RNA obtained from 5,000-10
000 microdissected preneoplastic cells
histologically determined to be prostatic interepithelial
neoplasia 2 (PIN2) cells. Double-stranded cDNA was
ligated to EcoRI adaptors, 5 cycles of PCR applied to the
cDNA with an adaptor-specific primer, and the resulting
PCR product subcloned into pAMP10 by the UBE-cloning
method (Life Technologies). Average insert size is 600
bp. NOTE: Not directionally cloned. This library was
constructed by David Krizman."

BASE COUNT 54 a 25 c 23 g 41 t
ORIGIN
Query Match 90.6%; Score 15.4; DB 10; Length 143;
Best Local Similarity 76.5%; Pred. No. 6e+02;
Matches 13; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTTCAGCATGATGATC 17
Db 52 TTTCAGCATGATGATC 36

RESULT 2
LOCUS BB318243 264 bp mRNA EST 11-JUL-2000
DEFINITION BB318243 RIKEN full-length enriched, adult male corpora
quadrigenima Mus musculus cDNA clone B2303/3p10 3', mRNA sequence.
ACCESSION BB318243
VERSION BB318243.1 GI:9025278
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 264)
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Arkawa, T., Carninci,
P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,
Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,
Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,
Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurinara, C., Kusakabe, M.,
Matsuyama, T., Miki, R., Mizuno, I., Nakamura, M., Oda, H., Okazaki, Y.,
Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata,
Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugihara, Y.,
Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomioka, N., Toy,
T., Tsunoda, Y., Watabiki, A., Watanabe, S., Yamamura, T., Yamataka, I.,
Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino,
M., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Kono, H., et al.)
Unpublished (2000)
Contact: Yoshihide Hayashizaki
Genome Exploration Research Group, Life Science Tsukuba Center,
Genome Science Laboratory
The Institute of Physical and Chemical Research (RIKEN), Genomic
Sciences Center
3-1-1 Koyada, Tsukuba, Ibaraki 305-0074, Japan
Tel: +81-298-36-9013
Fax: +81-298-36-9098
Email: genome-res@ric.riken.go.jp,
URL: <http://genome.riken.go.jp/>
Carninci, P., Matsuyama, Y., Watanabe, A., Itoh, M., Nagakawa, S., Sasaki,
N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Thermolabile and thermoelectrolytic of thermolabile enzymes by
thermostable and its application for the synthesis of full length
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2): 520-524 (1998)
Itoh, M., Katsunari, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,

Query Match	90.6%;	Score 15.4;	DB 156;	Length 335;
Best Local Similarity	76.5%;	Pred. No. 6.9e+02;		
Matches 13; Conservative	4;	Mismatches 0;	Indels 0;	Gaps, 0;

AUTHORS NIH-MGC <http://mgc.ncbi.nlm.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC).
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
 Tel.: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Eco RI site shown at the beginning of the sequence.
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.

constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTCTGCTCTTAAAGCTGG and 3' end
primer CGACCTGACGCTGACGACA."

BASE COUNT 132 a 97 c 83 g 162 t

ORIGIN

Query Match 90.6%; Score 15.4; DB 16; Length 474;
Best Local Similarity 76.5%; Pred. No. 7.2e+02;
Matches 13; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTYGARGARATGATCC 17
11:11:11:11:11:11
Db 364 TTTGAGGAATGATCC 348

RESULT 10
BG076302/c 499 bp mRNA EST 26-JAN-2001
LOCUS H3158B03-3 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
DEFINITION H3158B03 3', mRNA sequence.
ACCESSION BG076302 GI:12558871
VERSION BG076302.1
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 499)
AUTHORS Kargul, G., Dudekula, D. B., Qian, Y., Lim, M. K., Jaradat, S. A., Tanaka
T. S., Carter, M. G. and Ko, M. S. H.
TITLE Verification and initial annotation of NIA mouse 15K cDNA clone set
JOURNAL Unpublished (2001)
COMMENT Other ESTs: H3158B03-5
Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdaelg@sun-grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit <http://lgsun-grc.nia.nih.gov/cDNA/15K.html> for details.
Plate: H3158 row: B column: 03
Seq primer: -21M13 Forward
High quality sequence stop: 499
POLYA-No.

FEATURES
Source
1..499
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_id="H3158B03"
/clone_lib="NIA Mouse 15K cDNA Clone Set"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT) primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U.S.A., 97: 9127-9132; (2) Large scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of

7.5-day mouse embryos reveals enrichment in the t-complex
and under-representation on the X chromosome, 1998, Hum
Mol Genet 7: 1967-1978.

BASE COUNT 131 a 107 c 96 g 165 t

ORIGIN

Query Match 90.6%; Score 15.4; DB 173; Length 499;
Best Local Similarity 76.5%; Pred. No. 7.3e+02;
Matches 13; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTYGARGARATGATCC 17
11:11:11:11:11:11
Db 243 TTTGAGGAATGATCC 227

RESULT 11
BF453847 504 bp mRNA EST 01-DEC-2000
LOCUS maa/2e04.y1 Soares mouse 3NDMS Mus musculus cDNA clone
DEFINITION IMAGE:3822415 5' similar to SW:11L18.MOUSE P70380 INTERLEUKIN-18
PRECUSOR ; mRNA sequence.
ACCESSION BF453847 GI:11520016
VERSION BF453847
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 504)
AUTHORS NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:1458767
Seq primer: -40RP from Gibco
High quality sequence stop: 468.

FEATURES
Source
1..504
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_id="IMAGE:3822415"
/clone_lib="Soares mouse 3NDMS"
/sex="male"
/tissue_type="Spleen"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer (5'
GGTTCAGATCTGATGAGGAGCGCGGCTGCTGTTTCTTTTCTTTT
3'); double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. RNA
provided by Dr. Bertrand Jordan. Library went through
three rounds of normalization, and was constructed by
Bento Soares and M. Fatima Bonaldo."

BASE COUNT 164 a 96 c 105 g 138 t 1 others

ORIGIN

Query Match 90.6%; Score 15.4; DB 148; Length 504;
Best Local Similarity 76.5%; Pred. No. 7.3e+02;
Matches 13; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTYGARGARATGATCC 17
11:11:11:11:11:11

DB 147 TTGAGGAATGATCC 163

RESULT 12
LOCUS AM125102/c
DEFINITION AM125102 512 bp mRNA EST 22-OCT-1999
UI-M-BH2.1-apx-g-10-0-UI-S1 NIH_BMAP_M.S3.1 Mus musculus cDNA clone
AM125102 UI-M-BH2.1-apx-g-10-0-UI 3', mRNA sequence.
ACCESSION AM125102 GI:6100632
VERSION EST.
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 512)
TITLE Ronaldo, M.F., Lennon, G. and Soares, M.B.
Normalisation and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: MEST@mail.nih.gov

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized basal ganglia library cDNA library preparation: M.B. Soares lab clone distribution: NIH BMAP cDNA clones will be made available by the means that is soon to be determined. When NIH determines the means for distribution of the BMAP cDNA clones, this record will be updated accordingly when that means is determined. The following repetitive elements were found in this cDNA sequence: 17-81, >MSTP#LTR/MALR
Seg primer: M13 Forward
POLYA=Yes.

FEATURES

Source

Location/Qualifiers
1. 512
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BH2.1-apx-g-10-0-UI"
/clone_lib="NIH_BMAP_M.S3.1"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pTZ193-Pac (Pharmacia) with a modified polylinker. Site 1: Not I; Site 2: Eco RI; The NIH_BMAP_M.S3.1 library is a subtracted library of a series, ultimately derived from a mixture of individually tagged, normalized libraries from ten regions of the mouse brain (cerebellum, brain stems, olfactory bulbs, hypothalamus, cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus) after a series of subtractions to reduce the representation of cDNAs from which ESTs had already been generated. The following serially subtracted libraries were generated in this process: NIH_BMAP_M.S3.1, NIH_BMAP_M.S2, NIH_BMAP_M.S1. The subtracted library (NIH_BMAP_M.S3.1) was constructed as follows: PCR amplified cDNA inserts from NIH_BMAP_M.S2 clones from which 3' ESTs had been derived was used as a driver in a hybridization with the NIH_BMAP_M.S2 library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the NIH_BMAP_M.S3.1 library. This procedure has been

BASE COUNT 135 a 104 c 94 g 179 t
ORIGIN
Query Match 90.6%; Score 15.4; DB 111; Length 512;
Best Local Similarity 76.5%; Pred. No. 7.3e+02;
Matches 13; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGAGGARATGATCC 17
DB 374 TTGAGGAATGATCC 358

RESULT 13
LOCUS BG088827 564 bp mRNA EST 26-JAN-2001
DEFINITION H3158B03-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
H3158B03 5', mRNA sequence.
ACCESSION BG088827
VERSION BG088827.1 GI:12571391
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Kargul, G.J., Dudekula, D.B., Qian, Y., Lim, M.K., Jaradat, S.A., Tanaka
1 (bases 1 to 564)
TITLE Kargul, G.J., Dudekula, D.B., Qian, Y., Lim, M.K., Jaradat, S.A., Tanaka
Verification and initial annotation of NIA mouse 15K cDNA clone set
Unpublished (2001)
JOURNAL Other ESTs: H3158B03-3
COMMENT Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit <http://lgsun.grc.nia.nih.gov/cDNA/15k.html> for details.
Plate: H3158 row: B column: 03
Seg primer: -21M13 Reverse
High quality sequence stop: 564
POLYA=No.

FEATURES

Source

Location/Qualifiers
1. 564
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="H3158B03"
/clone_lib="NIA Mouse 15K cDNA Clone Set"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"
/note="Vector: pSPORT1, Site 1: SalI; Site 2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos , and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of

Previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)
TAG_LIB=NIH_BMAP_M.S3.1
TAG_TISSUE=basal_ganglia
TAG_SEQ=GTGAC"

REFERENCE 1 (bases 1 to 713)
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
TITLE Bonaldo,M.F., Lennon,G. and Soares,M.B.
JOURNAL Normalization and subtraction: two approaches to facilitate gene
MEDLINE genome Res. 6 (9), 791-806 (1996)
COMMENT 97044477
Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mestr@mail.nih.gov

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the basal ganglia tissue cDNA library preparation. M.B. Soares lab Clone distribution: Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMAP cDNAs whose availability will be considered under appropriate and limited collaborative arrangements. The following repetitive elements were found in this cDNA sequence: 17-81, >MSMD1TFR/MaLR
Seq primer: M13 Forward
POLYA=yes.

FEATURES

source

Location/Qualifiers
1..713
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U1-M-BG2-bcp-h-09-0-UT"
/clone_1lb="NIH_BMAP_MSC_S1"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pRT73D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; The NIH_BMAP_MSC_S1 library is a subtracted library derived from NIH_BMAP_MSC_N. NIH_BMAP_MSC_N was made from mouse spinal cord tissue. For a detailed description of the library from which this clone was derived, please visit our web site at brainest.eng.uiowa.edu.
TAG_LIB="NIH_BMAP_MSC_S1"
TAG_TISSUE="basal-ganglia"
TAG_SEQ="TGTCAC"
BASE COUNT 189 a 144 c 135 g 245 t
ORIGIN

Query Match 90.6%; Score 15.4; DB 142; Length 713;
Best Local Similarity 76.5%; Pred. No. 7.7e+02;
Matches 13; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTYGARGARATGAYCC 17
||:||||:||||:||||
Db 374 TTTGAGGAAATGATCC 358

RESULT 17
BE544528/c 719 bp mRNA EST 09-AUG-2000
LOCUS 601078354F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:3464221 5',
DEFINITION mRNA sequence.
ACCESSION BE544528
VERSION BE544528.1 GI:9773173
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 719)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: AFCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM8464 row: 1 column: 14
High quality sequence stop: 518.

FEATURES

source

Location/Qualifiers
1..719
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3464221"
/clone_1lb="NIH_MGC_12"
/tissue_type="Cervical carcinoma cell line"
/lab_host="DH10B"
/note="Organ: cervix; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.4 kb. Library prepared by Life Technologies."
BASE COUNT 229 a 155 c 148 g 187 t
ORIGIN

Query Match 90.6%; Score 15.4; DB 136; Length 719;
Best Local Similarity 76.5%; Pred. No. 7.7e+02;
Matches 13; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTYGARGARATGAYCC 17
||:||||:||||:||||
Db 321 TTTGAGGAAATGATCC 305

RESULT 18
AG007422/c 740 bp DNA GSS 07-FEB-1999
LOCUS AG007422
DEFINITION Homo sapiens genomic DNA, 21q region, clone: T1539X36, genomic survey sequence.
ACCESSION AG007422
VERSION AG007422.1 GI:3097375
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone: T1539X36.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 740)
AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (29-APR-1998) to the DDBJ/EMBL/GenBank databases.
Masanira Hattori, Kitasato University, Department of Science, JST Sequencing Laboratory; Kitasato 1-15-1, Sagamihara 228, Japan (E-mail:hattori@ngc.ims.u-tokyo.ac.jp, Tel:0427-78-9732, Fax:0427-78-9561)
2 (bases 1 to 740)
Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
Homo sapiens genomic DNA, chromosome 21q
Published Only in Database (1998) in press
Location/Qualifiers
1..740
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/clone="T1539X36"
/map="21q"

```

BASE COUNT      182 a      168 c      167 g      208 t      15 others
ORIGIN

Query Match      90.6%; Score 15.4; DB 229; Length 740;
Best Local Similarity 76.5%; Pred. No. 7.7e+02;
Matches 13; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTTGAGGAGATGATCC 17
        11:11:11:11:11:11
DB      285 TTTGAGGAGATGATCC 269

RESULT 19
BF793924/c      772 bp      mRNA      EST      12-JAN-2001
LOCUS      602254534f1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4346816 5',
DEFINITION      mRNA sequence.
ACCESSION      BF793924
VERSION      BF793924.1 GI:12098978
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 772)
AUTHORS      NIH-MGC http://mgc.ncl.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert.Strausberg@nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LAM9696 row: c column: 09
            High quality sequence stop: 710.
            Location/Qualifiers
                1..772
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_image="4346816"
                /clone_id="NIH_MGC_84"
                /tissue_type="adrenal cortex carcinoma, cell line"
                /lab_host="DH10B (phage-resistant)"
                /note="Organ: adrenal gland; Vector: PCMV-SPOrt6; Site_1:
                NotI; Site_2: SalI; Cloned unidirectionally; oligo-dt
                primed. Average insert size 1.229 kb. Library enriched for
                full-length clones and constructed by Life Technologies.
                Note: This is a NIH-MGC Library."

BASE COUNT      265 a      155 c      149 g      203 t
ORIGIN

Query Match      90.6%; Score 15.4; DB 169; Length 772;
Best Local Similarity 76.5%; Pred. No. 7.8e+02;
Matches 13; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTTGAGGAGATGATCC 17
        11:11:11:11:11:11
DB      172 TTTGAGGAGATGATCC 156

RESULT 20
BF542842/c      779 bp      mRNA      EST      09-AUG-2000
LOCUS      601067660f1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:3453758 5',
DEFINITION      mRNA sequence.
ACCESSION      BF542842

```

```

VERSION      BE542842.1 GI:9771487
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 779)
AUTHORS      NIH-MGC http://mgc.ncl.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert.Strausberg@nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: Incyte Genomics, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LAM8437 row: h column: 15
            High quality sequence stop: 657.
            Location/Qualifiers
                1..779
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_image="3453758"
                /clone_id="NIH_MGC_10"
                /cell_line="MGC36"
                /lab_host="DH10B"
                /note="Organ: Cervix; Vector: PCMV-SPOrt6; Site_1: NotI;
                Site_2: SalI; Cloned unidirectionally. Primer: oligo dt.
                Average insert size 1.5 kb. Library prepared by Life
                Technologies."

BASE COUNT      245 a      170 c      160 g      204 t
ORIGIN

Query Match      90.6%; Score 15.4; DB 136; Length 779;
Best Local Similarity 76.5%; Pred. No. 7.8e+02;
Matches 13; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTTGAGGAGATGATCC 17
        11:11:11:11:11:11
DB      134 TTTGAGGAGATGATCC 118

RESULT 21
A2530862/c      872 bp      DNA      GSS      03-NOV-2000
LOCUS      ENTAC19TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
DEFINITION      genomic DNA sequence.
ACCESSION      A2530862
VERSION      A2530862.1 GI:11085052
KEYWORDS      GSS.
SOURCE      Entamoeba histolytica.
ORGANISM      Entamoeba histolytica
Eukaryota; Entamoebidae; Entamoeba.
REFERENCE      1 (bases 1 to 872)
AUTHORS      Loftus, B., Van Aken, S., and Fraser, C.
TITLE      Determination of clone end sequences from Entamoeba histolytica
JOURNAL      HMI:IMSS sheared DNA library
COMMENT      unpublished (2000)
            Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: bjoftus@isig.org
            Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
            DNA library
            Seq primer: M13-Reverse

```

Class: shotgun
High quality sequence start: 62
High quality sequence stop: 727.
Location/Qualifiers
1. .872

FEATURES
source
/organism="Entamoeba histolytica"
/strain="HM1:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pROSI; Site: 1; Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999)."

BASE COUNT
316 a 120 c 153 g 277 t 6 others

ORIGIN

Query Match
Best Local Similarity 90.6%; Score 15.4; DB 223; Length 872;
Matches 13; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTYGARGARATGAYCC 17
||:||||:||||:|
Db 783 TTTGAAGAATGATGCC 767

RESULT 22
LOCUS CENS0447P 921 bp DNA GSS 18-MAY-2000
DEFINITION Tetradon nigroviridis genome survey sequence T7 end of clone 081E03 of library G from Tetradon nigroviridis, genomic survey sequence.

ACCESSION AL273742
VERSION AL273742.1 GI:7996005
KEYWORDS GSS; genome survey sequence.
SOURCE Tetradon nigroviridis.
ORGANISM Tetradon nigroviridis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Eurypterygii; Ctenosquamata; Acanthomorpha; Eucanthomorpha; Holacanthopterygii; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.
1 (bases 1 to 921)
Roest-Crolius,H., Jallion,O., Dasilva,C., Fizesma,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.

TITLE
JOURNAL
REFERENCE
AUTHORS
Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetradon nigroviridis
2 (bases 1 to 921)
Roest-Crolius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizesma,C., Wincker,P., Brothier,P., Quetier,F., Saurin,W. and Weissenbach,J.
Human gene number estimate provided by genome wide analysis using Tetradon nigroviridis DNA sequence
Unpublished

TITLE
JOURNAL
REFERENCE
AUTHORS
COMMENT
Direct Submission
Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetradon nigroviridis genome. For more information, please take a look at
<http://www.genoscope.cns.fr/tetradon>.

FEATURES
source
Location/Qualifiers
1. .921

BASE COUNT
224 a 199 c 220 g 273 t 5 others

ORIGIN

Query Match
Best Local Similarity 90.6%; Score 15.4; DB 231; Length 921;
Matches 13; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTYGARGARATGAYCC 17
||:||||:||||:|
Db 682 TTTGAGAGATGATGCC 666

RESULT 23
LOCUS BE544857/C 937 bp mRNA EST 09-AUG-2000
DEFINITION 601078856F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:3464700 5', mRNA sequence.

ACCESSION BE544857
VERSION BE544857.1 GI:9773502
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN at:
<http://image.jnl.gov>
Plate: LHM8465 row: P column: 13
High quality sequence stop: 683.
Location/Qualifiers
1. .937

FEATURES
source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3464700"
/clone_lib="NIH_MGC_12"
/tissue_type="cervical carcinoma cell line"
/note="Organ: cervix; Vector: PCMV-SPOrt6; Site: 1; NotI; Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.4 kb. Library prepared by Life Technologies."

BASE COUNT
306 a 202 c 196 g 233 t

ORIGIN

Query Match
Best Local Similarity 90.6%; Score 15.4; DB 136; Length 937;
Matches 13; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTYGARGARATGAYCC 17
||:||||:||||:|
Db 637 TTTGAGAGATGATGCC 621

[illegible]

	KEYWORDS	EST, house mouse.
SOURCE	Mus musculus	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus- I (bases 1 to 267)	
REFERENCE	Kondo,H., Aizawa,K., Akchira,S., Akiyama,J., Carninci,P., Endo,T., Fukuwa,S., Fukunishi,Y., Harai,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawaji,T., Kikuchi,N., Koijima,Y., Koya,S., Kusabe,H., Oka,H., Okazaki,Y., Owa,C., Ozawa,T., Saito,H., Sanjo,W., Satou,R., Shibata,K., Shibata,Suzuki,H., Takahashi,F., Tateno,M., Tomimaga,N., Tsunoda,Y., Wataniki,A., Wetanade,S., Yamamura,T., Yasunishi,A., Yokota,T., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y. RIKEN Mouse ESTs (Kondo,H., et al. 1999) Unpublished (1999) Contact: Yoshihide Hayashizaki Genome Exploration Research Group, Life Science Tsukuba Center, Genome Science Laboratory The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan Tel.: +81-298-36-9013 Fax: +81-298-36-9098 Email: genome-res@rictc.riken.go.jp/ URL:http://genome.ric.riken.go.jp/ Sasaki,N., Izawa,M., Matalhiki,M., Ozawa,K., Tanaka,T., Yoneda,Y., Matsura,S., Carninci,P., Muramatsu,M., Okazaki.Y. and Hayashizaki,Y.	
TITLE	JOURNAL	Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
COMMENT		Itoh.M., Kitsumagi.T., Akiyama.J., Shibata.K., Izawa.M., Kawasaki.I.,Tomaru.Y., Carninci.P., Shibata.Y., Ozawa.Y., Muramatsu.M., Okazaki.Y. and Hayashizaki.Y. Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5) : 463-470 (1999) Carninci P. and Hayashizaki Y. High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999) Further details: Location/Qualifiers 1..267 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone_abbrev="A930422M09" /clone_idb="RIKEN full-length enriched, adult male testis (DH10B)" "/sex="male" /library_type="testis" /dev_stage="adult" /_lib_host="DH10B" /_note="Site_1: SalI; Site_2: BamHI; CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer 15' GAGACGGAGAGATCCCAAGACTCTTTTTCCTTTTTTTTNNN 3'}. cDNA was prepared by using reverse transcriptase therm-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence 15' GAAGAAGATTCTCGATTAAATAATTAATGCCCCCCCCC 3'. cDNA was cloned into the XhoI and BamH1 sites. Vector: a modified phagescript KS(+) after bulk excision from lambdaPhi C12Cloning sites, 5' end: SalI, 3' end: BamH1."
BASE COUNT		51 a 69 c 75 g 72 t
ORIGIN		

Query Match 84.7%; Score 14.4; DB 26; Length 267;
Best Local Similarity 75.0%; Pred. No. 2.3e+03;
Matches 12; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTYGARGARATGAYC 16
||:||||:||||:|
DB 13 TTTGAGGAGATGACC 28

Search completed: May 15, 2001, 10:54:39
Job time: 2945 sec